AD-A269 668

(102)

University of Southern California

Knowledge resources or decisions?

Cecile L. Paris and Elisabeth A. Maier
USC/ Information Sciences Institute
4676 Admiralty Way
Marina del Rey, California 90292

August 1991
ISI/RR-93-322

DTIC
ELECTE
SEP 22 1993
S E D

[redacted] STATEMENT A
Approved for public release
Distribution Unlimited

INFORMATION SCIENCES INSTITUTE

310/822-1511
4676 Admiralty Way/Marina del Rey/California 90292-6695

93-22210

Knowledge resources or decisions?

Cecile L. Paris and Elisabeth A. Maier
USC/ Information Sciences Institute
4676 Admiralty Way
Marina del Rey, California 90292

August 1991
ISI/RR-93-322

Accesion For
NTIS CRA&I [X]
DTIC TAB [ ]
Unannounced [ ]
Justification .....
By .....
Distribution /
Availability Codes
Dist | Avail and / or Special
A-1 | 20

DTIC ELECTE SEP 22 1993 S E D

In IJCAI-91 Workshop on Decision Making throughtout the Generation Process, pages 11-17

DTIC ...

93-21890

STATEMENT A
Approved for public release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

FORM APPROVED
OMB NO. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching exiting data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimated or any other aspect of this collection of information, including suggestings for reducing this burden to Washington Headquarters Services, Directorate for information Operations and Reports, 1215 Jefferson Davis highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	August 1991	Research Report

4. TITLE AND SUBTITLE
Knowledge Resources or Decisions?

5. FUNDING NUMBERS
NCC2-520
DABT63-91-C-0025

6. AUTHOR(S)
Cecile L. Paris and Elisabeth A. Maier

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
USC INFORMATION SCIENCES INSTITUTE
4676 ADMIRALTY WAY
MARINA DEL REY, CA 90292-6695

8. PERFORMING ORGANIZATON REPORT NUMBER
RR-322

9. SPONSORING/MONITORING AGENCY NAMES(S) AND ADDRESS(ES)
NASA AMES
Moffett Field, CA 94035

ARPA
3701 Fairfax Drive
Arlington, VA 22203

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
In IJCAI-91 workshop on Decision Making throughout the Generation Process, pages 11-17

12A. DISTRIBUTION/AVAILABILITY STATEMENT
UNCLASSIFIED/UNLIMITED

12B. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

In this paper we argue that the problem of decisions can only be discussed when the resources which contribute to the process of text generation are identified. We claim that declarative and procedural knowledge-while resources correspond to the former decisions to the latter - have to be clearly separated. After evaluating various text planning systems from this angle we outline what the consequences for the design of a text generation system are.

14. SUBJECT TERMS
natural language generation, text planning system, resources, decisions

15. NUMBER OF PAGES
6

16. PRICE CODE

17. SECURITY CLASSIFICTION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

GENERAL INSTRUCTIONS FOR COMPLETING SF 298

The Report Documentation Page (RDP) is used in announcing and cataloging reoprts. It is important that this information be consistent with the rest of the report, particularly the cover and title page. Instructions for filling in each block of the form follow. It is important to stay within the lines to meet optical scanning requirements.

Block 1. Agency Use Only (Leave blank).

Block 2. Report Date. Full publication date including day, month,a nd year, if available (e.g. 1 jan 88). Must cite at least the year.

Block 3. Type of Report and Dates Covered. State whether report is interim, final, etc. If applicable, enter inclusive report dates (e.g. 10 Jun 87 - 30 Jun 88).

Block 4. Title and Subtitle. A title is taken from the part of the report that provides the most meaningful and complete information. When a report is prepared in more than one volume, repeat the primary title, add volume number, and include subtitle for the specific volume. On classified documents enter the title classification in parentheses.

Block 5. Funding Numbers. To include contract and grant numbers; may include program element numbers(s), project number(s), task number(s), and work unit number(s). Use the following labels:

C - Contract	PR - Project
G - Grant	TA - Task
PE - Program Element	WU - Work Unit Accession No.

Block 6. Author(s). Name(s) of person(s) responsible for writing the report, performing the research, or credited with the content of the report. If editor or compiler, this should follow the name(s).

Block 7. Performing Organization Name(s) and Address(es). Self-explanatory.

Block 8. Performing Organization Report Number. Enter the unique alphanumeric report number(s) assigned by the organization performing the repor.

Block 9. Sponsoring/Monitoring Agency Names(s) and Address(es). Self-explanatory

Block 10. Sponsoring/Monitoring Agency Report Number. (If known)

Block 11. Supplementary Notes. Enter information not included elsewhere such as: Prepared in cooperation with...; Trans. of ...; To be published in... When a report is revised, include a statement whether the new report supersedes or supplements the older report.

Block 12a. Distribution/Availability Statement. Denotes public availability or limitations. Cite any availability to the public. Enter additional limitations or special markings in all capitals (e.g. NOFORN, REL, ITAR).

DOD	- See DoDD 5230.24, "Distribution Statements on Technical Documents."
DOE	- See authorities.
NASA	- See Handbook NHB 2200.2.
NTIS	- Leave blank.

Block 12b. Distribution Code.

DOD	- Leave blank.
DOE	- Enter DOE distribution categories from the Standard Distribution for Unclassified Scientific and Technical Reports.
NASA	- Leave blank.
NTIS	- Leave blank.

Block 13. Abstract. Include a brief (Maximum 200 words) factual summary of the most significant information contained in the report.

Block 14. Subject Terms. Keywords or phrases identifying major subjects in the report.

Block 15. Number of Pages. Enter the total number of pages.

Block 16. Price Code. Enter appropriate price code (NTIS only).

Blocks 17.-19. Security Classifications. Self-explanatory. Enter U.S. Security Classification in accordance with U.S. Security Regulations (i.e., UNCLASSIFIED). If form contins classified information, stamp classification on the top and bottom of the page.

Block 20. Limitation of Abstract. This block must be completed to assign a limitation to the abstract. Enter either UL (unlimited) or SAR (same as report). An entry in this block is necessary if the abstract is to be limited. If blank, the abstract is assumed to be unlimited.

Knowledge resources or decisions?

Cécile L. Paris*
Information Sciences Institute
of the University of Southern California
4676 Admiralty Way
Marina del Rey, CA 90292-6695
U.S.A.
paris@isi.edu

Elisabeth A. Maier†
Projekt KOMET
GMD-IPSI
Dolivostr. 15
D-W-6100 Darmstadt
Germany
maier@ipsi.darmstadt.gmd.de

Abstract

In this paper we argue that the problem of *decisions* can only be discussed when the *resources* which contribute to the process of text generation are identified. We claim that declarative and procedural knowledge - while resources correspond to the former and decisions to the latter - have to be clearly separated. After evaluating various text planning systems from this angle we outline what the consequences for the design of a text generation system are.

1 Introduction

In this paper, we argue that *all* the linguistic knowledge necessary to construct a text should be declaratively represented as *resources* available to a text planner.

Many systems which work on the level of sentence generation represent their resources in a declarative form. The grammar, for example, can be considered as a declarative resource. Further, semantic information used in sentence generation has often been represented declaratively in a knowledge base.

We want to extend this view for the generation of texts rather than sentences and therefore need to specify clearly for the process of text generation (and text planning, respectively) what defines resources and decisions.

We claim that without a proper declarativization of the resources used in text generation, it becomes difficult to distinguish *decisions* made by the system from constraints imposed by the resources of language. We define a decision to be a choice made by the system where the information provided by the resource is not sufficient to make one distinct choice. The resources must be clearly defined before appropriate decisions can be identified. Failure to do so runs the danger of requiring decisions when none are justifiable or appropriate.

In many current approaches to text planning systems, although some of the linguistic knowledge is represented explicitly, much of it is still implicitly encoded in the planner itself and gives rise to false and unnecessary decisions, which would not occur were this information to be explicitly represented outside the control knowledge of the planner. We believe this is wrong and that, before we can even discuss the decisions that take place in text generation, we have to strive to represent all linguistic information as resources, declaratively and separately from the control information. We also have to represent this information in a modular way, grouping together information responsible for the same phenomena. To take these design criteria into account is important for the same reasons raised in both general debates in computer science regarding issues of "declarativity *vs* procedurality", and more recent essential considerations of modularity, modifiability and generalization of knowledge.

What we therefore aim at is to draw a clear borderline between linguistic knowledge to be encoded declaratively and operational knowledge which controls the flow of communication in and between the modules. Getting the resource/decision distinction right means that:

- Modularity and transparency is enhanced. One can easily examine the resources available. Furthermore, when knowledge is implicitly encoded in the decisions made by a system, there is a danger in mixing various kinds of knowledge in one complex decision.
- Modifiability and extendability is increased: because of the modularity and declarativity, it is easy to make changes and extend the system.
- Generalizability is also increased: when knowledge is represented declaratively, it becomes easier to notice and represent generalizations.

Getting this distinction wrong almost guarantees unexpandable (and thus short lived) systems.

2 Resources and decisions

In generation, resources represent linguistic potential available to construct a text. This is, they contain all the knowledge necessary to produce the linguistic phenomena specifically needed for the generation of text. At the lexicogrammatical level, these resources include at least a grammar and a lexicon. At the text planning level, it is still unclear exactly what these resources are and which information they cover. Text types, theme

*Her research is supported by DARPA under grant MDA 903-81-C-0335 and under NASA Ames cooperative agreement number NCC 2-520.

†Her work at ISI is supported by RADC grant FQ7619-89-03326-0001.

development, rhetorical relations, etc. are types of information that come to mind. Necessarily, none of the resources can be seen as independent of all of the others. Therefore, the *interaction* among resources and the way they influence each other is also *part* of the resources. All the resources comprise what is available to the text planner to generate a text. We should now strive to identify all the resources needed for the generation of text and represent them explicitly and in a modular way.

As the linguistic potential expressed in the resources might under-determine the solution space, there will be a set of alternatives for satisfying the initial specification as opposed to one solution. At this point, there is a need to decide on the alternative to be preferred. *Decisions* occur during this phase.

It is a weakness of many current approaches to text generation that the set of alternatives is not sufficiently restricted by the resources themselves, thus rendering an eventual decision more complex than need to be. We believe this is so because, in many cases, all the linguistic information has not been clearly identified and represented in an explicit and modular way. This is especially the case at the text organization level, where it is still rather vague as to what the resources and their interactions are and how one can represent them declaratively.

In recent computational linguistic work, there has already been a push towards representing linguistic knowledge at the level of the grammar and the lexicon in a declarative and modular way, separating different concerns, e.g, (Emele *et al.*, 1990; Emele *et al.*, 1991). We are now arguing that we should also try to represent *all* the linguistic resources needed for the generation of *text* in a declarative way, thus also enhancing their modularity, modifiability, and generality.

3 Identification of information required throughout the text generation process

In the following we discuss the various types of information required in the text planning process. We describe how this information was represented in recent text planning systems and show how linguistic information was encoded both in procedural and declarative form. We then give a short outline of the subtasks in text generation which are to be supported by declarative knowledge sources.

3.1 Linguistic and control knowledge embodied in text planning systems

Different concerns have often been mixed together in text planning systems, resulting in more complex and opaque systems, as well as systems that eventually become harder to augment and maintain. Linguistic information and control knowledge in particular have often been mixed, and much of what is considered control knowledge should be more appropriately treated as part of the declarative resource representation as linguistic information. The decision process has thus been complicated by not representing separately, declaratively, and in a modular way all the linguistic information that needs to be available to the text planner.

There are indeed numerous examples of these phenomena, of which we only mention a few here:

- In systems such as TEXT (McKeown, 1985) or TAILOR (Paris, 1988), a rhetorical strategy is chosen by making a decision based on the communicative goal given to the text planner and the knowledge available in the knowledge base. A better solution is to express the set of potential communicative goals and the rhetorical strategies that can be used to achieve these goals as declarative resources available to the text planner. The interactions between domain knowledge, communicative goals and rhetorical strategies should be stated explicitly and in a modular way. What would previously have been represented as an extrinsic 'decision' then follows as an automatic consequence of the definition of the resources themselves. In TAILOR, an additional constraint is added: the user model; again, if the choice of strategy is consistent for a given type of user model, this linguistic information can be represented separately, instead of being a decision of the planner (Paris and Bateman, 1990).

- Focus handling: in many planners, focus rules are embedded in the code of the text planner. They are not represented as a linguistic resource available to the text planner.[1]

- Constraints on the style of text to be generated: in systems such as Pauline (Hovy, 1988a), planning and realization decisions are made on stylistic grounds by appealing to so-called Rhetorical Goals, which are distinguished from higher-level pragmatic (interpersonal and situational) concerns. While these Rhetorical Goals are an attempt both to modularize stylistic issues and to represent in declarative form the kinds of knowledge needed to generate pragmatically/stylistically appropriate text, Pauline did not go nearly far enough: the system's calls to the Rhetorical Goals for guidance at decision points are all procedurally implemented and monitored.

- Selection heuristics based on previous discourse (Moore, 1989). To decide upon the specific plan to choose when several are available, some systems employ 'selection heuristics' that make a decision based on previous discourse, user model, or other pragmatic factors. These are embedded in the code. It would be more appropriate to represent them as additional resources to the text planner (Moore and Paris, 1991).

- In various approaches to text generation, a text is planned by posting a communicative goal and choosing a plan operator capable of achieving that goal - e.g., (Appelt, 1981; McKeown, 1985; Hovy,

[1] An attempt to make focus handling more explicit and integrate text structure planning with a separate level of control for focus appears in (Hovy and McCoy, 1989).

1988b; Moore and Paris, 1989). The operator in turn can post other subgoals to be achieved. Plan operators indicate when a specific rhetorical relation can be used given the knowledge base at hand, the communicative goals being pursed, and a user model; or how a communicative goal is to be decomposed into other subgoals, again depending on the knowledge available, a user model or rhetorical relations. They indeed represent in a declarative way some of the linguistic information available for constructing a text. However, some of the linguistic knowledge (such as focus handling or use of past discourse) is embedded in the control knowledge and is hard to examine, change and augment.

Furthermore, what is not represented in a modular way are the interactions between the various types of information: as mentioned above, a plan operator includes constraints on rhetorical relations, user model, and domain knowledge at hand. The number of plan operators increases as the range of produced text increases and as the domain of application expands. In such cases, there is a clear need to impose further organization. Indeed, as the number and the specificity of the planning operators increase, so does their complexity and the complexity of their interactions. A specific plan operator might be employed in the service of several higher level communicative goals; a plan operator might occur in many different text types, correspond to various types of knowledge in the knowledge base, etc. It is now necessary to reduce the complexity of plan operators by reducing the overall power which each operator has. Instead of mixing into one operator constraints on the knowledge base, the rhetorical relations, a user model, etc., we believe it is preferable to state instead the constraints and interactions among all these in an explicit and modular fashion (Moore and Paris, 1991).

Based on our experience and understanding of the generation process, it is time now to re-examine our text planners and identify where they mix linguistic information with decisions, where they have failed to represent linguistic information, giving rise to 'ill-founded' decisions, and where a more declarative and modular representation of the resources is necessary.

We must thus now look more precisely at the types of knowledge embodied in plan operators such as the ones mentioned above or in the various 'decisions' taken by generation systems and note where generalizations can be made. These form the basis for identifying linguistic resources required at the text organization level.

3.2 Linguistic phenomena as indicators for necessary knowledge resources

To determine the various types of information necessary for the development of declarative linguistic resources it is also necessary to closely examine the phenomena to be dealt within the text generation process. By doing this we may gain additional indications which information types are possible candidates for representation as a knowledge source. Among these are:

- text structure and the order of textual units:
  The system has to provide resources containing a set of possible text structuring means together with the description of their conditions of use. This component must also have a notion of how to order the elements of a text structure in order to increase ease of understanding.

- parataxis, hypotaxis and cohesion:
  Once the text structure is determined it still is unclear how the information is to be ordered with respect to its syntactic structure. The system must contain a collection of motivations which indicate under which circumstances parataxis is to be preferred over hypotaxis etc. A similar problem also to be addressed by a resource is the determination of sentence boarders, i.e. the question of how much information is to be given in one single sentence.

- thematic progression and focus shift:
  Other resources have to describe a potential of possibilities how to construct a coherent sequence of sentences and how these are linked by thematic progression.

- lexical choice:
  The choice of adequate lexical items is another step in the process of text generation which has to be supported by distinct knowledge sources. These resources not only are responsible for the selection of lexical items most adequate in the context given, they also have to be able to construct cohesive chains of lexemes.

- relevance of textual content:
  Other components have to deal with the determination of concepts relevant in a given situation. This component must also provide knowledge of how much information concerning a certain topic is to be selected.

Although it is not clear yet whether the generation of each of these phenomena corresponds to only one resource responsible for its treatment, it is obvious that each phenomenon should be defined with respect to the resources responsible.

Our main effort at this point should thus be to clearly identify and represent these resources. Only then can the issue of decision-making in generation be addressed.

In our own work, we believe that the breakdown of resources offered by systemic functional linguistics provides an appropriate framework to study and identify the linguistic resources needed at the text planning level, and we are investigating within this framework the resources necessary for planning a text.

4 Identifying the resources needed: a start

It has been shown that the organization of the lexicogrammatical resources proposed within Systemic Func-

tional Linguistics (SFL) (Halliday, 1978) is not only compatible with the goals of modularity and declarativity, as we have described them, but is also a good candidate for achieving these goals (Bateman and Momma, 1991).

SFL offers one breakdown of resources that we believe is useful to start studying and organizing the knowledge sources available to a text planner: the meta-functional split. SFL distinguishes three types of meanings occurring in language: *ideational, interpersonal* and *textual* meaning. While ideational meaning represents the world as it is and how it is experienced, interpersonal meaning is concerned with information about the participants of the discourse and their social relations. The textual metafunction deals with the presentation of both ideational and interpersonal information as text.

The meta-functional split was used as motivation for the development of an architecture for text generation systems as described in (Matthiessen, 1987; Matthiessen and Bateman, 1991). Corresponding to the three metafunctions, three modules or so-called *bases* are distinguished. The ideation base captures the experiential knowledge, the interaction base copes with interpersonal meaning and the text base is responsible for the treatment of textual phenomena. These three modules form the basis for a possible architecture for text generation systems. It is not enough, though, and one needs to specify what goes into each module as well as their interactions.

The ideation base has already been studied extensively for the last few years. It has been incorporated into the text generation system PENMAN (Mann, 1983) as a linguistically motivated hierarchy of concepts — the Upper Model (Bateman *et al.*, 1990; Bateman, 1990; Halliday and Matthiessen, in preparation) — which is augmented in specific applications by subordinating to it domain-specific knowledge — the Domain Model. The Upper Model is one of the resources which co-constrains the resources of the grammar according to ideational function.

The *text base*, that is, that area of the semantics that controls the *textual resources* of the grammar such as theme-rheme structures, the given-new distinction, etc., is still largely unstudied, although there has recently been a lot of research on discourse phenomena in text generation, as e.g. in (Dale, 1989) or (Lascarides and Asher, 1991). It is still rather vague as to what type of knowledge belongs in the text base and how it should be organized. And yet, it is necessary to clearly define these resources, especially for the development of a text planning system.

In order to define them, we can study information types from various angles, by identifying:

1. what information is necessary to control the lexicogrammatical resources of the generation system, as e.g. the grammar, the lexicon, etc.;

2. what is necessary to produce discourse.

The Upper Model, for example, was mainly derived by studying the ideation base from the first angle, i.e. from the point of view of sentence generation.

For the text base, some research has already been conducted taking the first viewpoint (Bateman and Matthiessen, to appear; Matthiessen and Bateman, 1991). In that work, Bateman and Matthiessen identified that the text base needs to include at least knowledge to control areas of the grammar such as conjunctions, thematizations, and voice choice. From this they argue that the text base must include information such as rhetorical relations, potential topical shifts, and levels of textual statuses (newsworthiness, thematicity, etc).

Studying the text base from the other angle, we must, as mentioned above, examine carefully the text planners we currently have and identify the linguistic information that is embodied within them but not represented as resources. We believe a lot will be gained by this enterprise, and it will allow us, finally, to start discussing the issue of decision-making in text generation.

5 The representation of knowledge using constraint-based formalisms

It is obvious that various knowledge sources cannot be discussed in isolation. This can be observed from the phenomena as discussed above: most of the phenomena depend on each other. As mentioned above, interactions among knowledge resources are part of the information sources. Thus the linguistic resources are to be defined with respect to each other. One way to do this is to express these linguistic resources in terms of constraints. They entail a deduction mechanism by which a partial specification of any kind can be filled out to give a set of maximally information-filled specifications consistent with all resources. In other words, given these resources, there is a need to have a mechanism to utilize them, in order to obtain the set of alternatives that satisfy an initial specification for a text. This mechanism ensures that constraints on the resources and their interactions are satisfied.

This is obviously still not enough. There is also a need for some control mechanism to make sure the constraint propagation is performed in some orderly fashion (e.g., to make it more efficient, or psychologically motivated, etc.).

By choosing one uniform constraint-based formalism for the representation of knowledge sources needed for the generation of *text* we also get rid of the border between the textual and the sentence or rather the lexicogrammatical level. The result is one uniform architecture for the generation of text, instead of distinguishing - as done so far in the history of text generation - between text planning and sentence generation. Conflicts occurring between text and sentence planning are avoided since the resources of both levels no are able to constrain each other.

Additionally we gain an advantage common to constraint-based mechanisms, which is the neutrality of the formalism to a certain direction of processing. This means that grammars represented in this way can be employed both for parsing and analysis. Therefore, textual knowledge formulated in a constraint-based formalism is also valuable for *text analysis*.

Finally, this formalism cannot be considered as only a

means to represent knowledge; it can also be used as a methodology to develop linguistic components. By forcing the developer to state knowledge declaratively the possibility to mix control information with linguistic knowledge no longer exists.

6 Conclusions

Our current understanding of the text planning process is one of mutual activation and constraint of the resources available to a text planner. These resources should embody all the linguistic information necessary to produce a text as well as the interactions among various types of information. It is a weakness of many current approaches that the set of alternatives is not sufficiently restricted by the resources, because much of the linguistic knowledge still resides in the control knowledge. This overburdens the decision-making. Identifying the knowledge resources and their interactions should now become a primary concern. Only when these have been identified and represented declaratively and in a modular way can we start discussing the decisions that must take place.

Acknowledgments

This direction of this paper is based on discussions of the ISI text planning group consisting of Eduard Hovy, Julia Lavid, Vibhu Mittal, Richard Whitney, and the authors. Giuseppe Carenini has also contributed to this lively exchange of ideas during his visit to ISI. This view, however, represent our particular slant on the discussions and may not be shared by all members of the group. Finally, John Bateman, our roving (and raving!) systemic representative, provided useful comments, inspiration, and criticisms on these premature ideas. We also want to thank Marie Meteer, Dave McDonald and Mick O'Donnell for comments on an earlier draft of this paper.

References

(Appelt, 1981) Doug E. Appelt. *Planning Natural Language Utterances to Satisfy Multiple Goals*. PhD thesis, Stanford University, Stanford, Ca, 1981. also a Technical report No 259, SRI Internatinal, Artificial Intelligence Center, 1982.

(Bateman and Matthiessen, to appear) John A. Bateman and Christian M.I.M. Matthiessen. Uncovering the text base. In Hermann Bluhme and Hao Keqi, editors, *Selected Papers from the International Conference on Research in Text and Language, Xi'an Jiaotong University, Xi'an, P.R. China, 29-31 March 1989*. to appear.

(Bateman and Momma, 1991) John A. Bateman and Stefan Momma. The nondirectional representation of Systemic Functional Grammars and Semantics as Typed Feature Structures. Technical report, GMD/Institut für Integrierte Publikations- und Informationssysteme, Darmstadt and Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, January 1991.

(Bateman *et al.*, 1990) John A. Bateman, Robert T. Kasper, Johanna D. Moore, and Richard A. Whitney. A general organization of knowledge for natural language processing: the PENMAN upper model. Technical report, USC/Information Sciences Institute, Marina del Rey, California, 1990.

(Bateman, 1990) John A. Bateman. Upper modeling: organizing knowledge for natural language processing. In *5th. International Workshop on Natural Language Generation, 3-6 June 1990*, Pittsburgh, PA., 1990. Organized by Kathleen R. McKeown (Columbia University), Johanna D. Moore (University of Pittsburgh) and Sergei Nirenburg (Carnegie Mellon University).

(Dale, 1989) Robert Dale. Cooking up referring expressions. In *Proceedings of the Twenty-Seventh Annual Meeting of the Association for Computational Linguistics*, Vancouver, British Columbia, June 1989. Association for Computational Linguistics.

(Emele *et al.*, 1990) Martin C. Emele, Ulrich Heid, Stefan Momma, and Rémi Zajac. Organizing linguistic knowledge for multilingual generation. In *13th. International Conference on Computational Linguistics (COLING-90)*, Helsinki, Finland, 1990. Also available as: Project Polygloss Paper, Institut für Maschinelle Sprachverarbeitung, University of Stuttgart, Germany.

(Emele *et al.*, 1991) Martin Emele, Ulrich Heid, Stefan Momma, and Rémi Zajac. Interactions between linguistic constraints: Procedural vs. declarative approaches. *Journal of Machine Translation*, 1991. Special edition on the role of text generation in MT.

(Halliday and Matthiessen, in preparation)
Michael A.K. Halliday and Christian M.I.M. Matthiessen. *The Bloomington Lattice*. de Gruyter, Berlin, in preparation.

(Halliday, 1978) Michael A. K. Halliday. *Language as social semiotic*. Edward Arnold, London, 1978.

(Hovy and McCoy, 1989) Eduard H. Hovy and Kathleen F. McCoy. Focusing your RST: A step towards generating coherent multisentential text. In *Proceedings of the 11th. Annual Conference of the Cognitive Science Society*, pages 667–674, University of Michigan, Ann Arbor, Michigan, August 16-19 1989. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

(Hovy, 1988a) Eduard H. Hovy. *Generating natural language under pragmatic constraints*. Lawrence Erlbaum, Hillsdale, New Jersey, 1988.

(Hovy, 1988b) Eduard H. Hovy. Planning coherent multisentential texts. In *The Proceedings of the 26th. Annual Meeting of the Association of Computational Linguistics*, pages 163 – 169. Association for Computational Linguistics, June 1988.

(Lascarides and Asher, 1991) Alex Lascarides and Nicholas Asher. Discourse relations and defeasible knowledge. In *Proceedings of the 1991 Annual Meeting of the Association for Computational Linguistics*, Berkeley, California, 1991. Association for Computational Linguistics.

(Mann, 1983) William Mann. An Overview of the Penman Text Generation System. Technical Report ISI/RR-83-114, USC/Information Sciences Institute, 1983.

(Matthiessen and Bateman, 1991)
Christian M.I.M. Matthiessen and John A. Bateman. *Text generation and systemic-functional linguistics: experiences from English and Japanese*. Frances Pinter, London, 1991.

(Matthiessen, 1987) Christian M.I.M. Matthiessen. Notes on the organization of the environment of a text generation grammar. In Gerard Kempen, editor, *Natural Language Generation: Recent Advances in Artificial Intelligence, Psychology, and Linguistics*. Kluwer Academic Publishers, Boston/Dordrecht, 1987. Paper presented at the Third International Workshop on Natural Language Generation, August 1986, Nijmegen, The Netherlands.

(McKeown, 1985) Kathleen R. McKeown. *Text Generation: Using Discourse Strategies and Focus Constraints to Generate Natural Language Text*. Cambridge University Press, Cambridge, England, 1985.

(Moore and Paris, 1989) Johanna D. Moore and Cécile L. Paris. Planning text for advisory dialogues. In *Proceedings of the Twenty-Seventh Annual Meeting of the Association for Computational Linguistics*, pages 203 – 211, Vancouver, British Columbia, June 1989.

(Moore and Paris, 1991) Johanna D. Moore and Cécile L. Paris. The EES text planner and its architecture. Presented at the AAAI workshop on *Explanation Architectures*, Anaheim, July 1991.

(Moore, 1989) Johanna D. Moore. *A Reactive Approach to Explanation*. PhD thesis, UCLA, December 1989.

(Paris and Bateman, 1990) Cécile L. Paris and John A. Bateman. User modeling and register theory: A congruence of concerns. ISI Technical Report, 1990.

(Paris, 1988) Cécile L. Paris. Tailoring Object Descriptions to the User's Level of Expertise. *Computational Linguistics*, 14 (3):64–78, September 1988. Special Issue on User Modeling.